

第 1 部分 辅导材料

第 1 章 多媒体计算机概述

1.1 本章要点

- (1) 多媒体计算机的定义、分类,以及多媒体计算机和普通计算机有什么不同,多媒体计算机要解决的关键技术。
- (2) 多媒体技术促进了通信、娱乐和计算机的融合。特别是多媒体技术是解决高清晰度电视切实可行的方案。用多媒体技术制作 DVD、影视音响卡拉 OK 机,以及多媒体家庭网关。
- (3) 多媒体计算机技术的应用和发展:多媒体数据库、多媒体通信和多媒体创作工具,以及多媒体计算机的发展趋势。

1.2 重点与难点内容分析

1.2.1 多媒体计算机的定义和分类

定义:计算机综合处理多媒体信息(文本、图形、图像、音频和视频)使多种信息建立逻辑链接,集成成为一个系统并具有交互性的技术。简单地说,计算机综合处理声、文、图信息,具有集成性和交互性。

从开发和生产厂商以及应用的角度出发,多媒体计算机可以分为两大类。

(1) 家电制造厂商研制的电视计算机(teleputer):把 CPU 放到家电中,通过编程控制管理电视机、音响。有人称它为“灵巧”电视(smart TV)。

(2) 计算机制造厂商研制的计算机电视(compuvision):采用微处理器(80x86)作为 CPU,其他设备还有 VGA 卡、CD-ROM、音响设备,以及扩展的多媒体家电系统。有人说它的发展方向是 TV-killer。

1.2.2 多媒体计算机需要解决的关键技术

(1) 视频音频信息的获取技术

获取视频信号的方法有以下 3 种:

① 利用计算机产生彩色图形、静态图像和动态图像。

② 利用彩色扫描仪,扫描输入彩色图形和静态图像。

③ 利用视频信号数字化仪把彩色全电视信号数字化后,输入到多媒体计算机中,获得静态和动态图像。

(2) 多媒体数据压缩编码和解码技术

多媒体数据压缩编码和解码技术是多媒体系统的关键技术。多媒体系统具有综合处理文字、图形、图像、动画、声音以及视频等信息的能力。为了得到满意的视听效果,要求处理大量的数字化声音和视频信息。由于声音和视频的信息量非常大,如果在未进行压缩的情况下实现动态视频及立体声的实时处理,对于目前的微型计算机是无法实现的,因此,必须对多媒体信息进行压缩编码和解压缩编码的处理。关于多媒体数据压缩已形成了许多标准,如静态图像压缩标准(Joint Photographic Experts Group, JPEG)。在压缩编码中还用到许多算法编码,如预测编码、变换编码、统计编码、混合编码等。

(3) 视频音频数据的实时处理和特技

图像信息一般是二维信号,如一幅图像通常由 512×512 个像素组成。每个像素有 256 级灰度,或者是 $3 \times 8b$ RGB(红、绿、蓝)的 16×2^{20} 种颜色。一幅图像就有 256KB 或 768KB(彩色图像)个数据。为了完成视觉处理的传感,要利用预处理、分割、扫描、识别和解释多种处理以及数学的运算、点处理、二维卷积运算、二维正交变换、坐标变换、统计量计算等。

(4) 视频音频数据的输出技术

通过扫描仪和视频信号获取器,静态的图像和运动视频信号可以数字化后存储到帧存储器中,现在要解决如何把它们变成标准文件存储到内存或外存,同时还需要解决如何将不同的图像文件格式进行转换,在显示器上输出。

1.2.3 多媒体技术促进了通信、娱乐和计算机的融合

(1) 多媒体技术是解决电视数字化及 HDTV 的可行方案

应用多媒体技术制造的高清晰度电视(HDTV)可以支持任意分辨率的输出,而且输入输出分辨率可以独立,输出分辨率也可以任意地改变。可以用任意的窗口尺寸输出,同时还具备许多新的功能,如图形功能、视频音频特技以及交互功能。

高清晰度数字电视技术及交互式电视技术采用了数字式视频、数字式音频及 MPEG 压缩编码算法以便于数据传输、存储、计算机控制和管理。国际标准 MPEG-II 提供了 4 种工具,即空间可扩展性、时间可扩充性、信噪比可扩充性及数据分块等。

(2) 利用多媒体技术制作 VCD、DVD 及影视音响

应用多媒体计算机技术可制作 VCD、DVD 影视音响卡拉 OK 机等。基于 ES-3204 芯片的 VCD 系统框图见主教材图 1.1。

VCD 播放机是从 CD-ROM 驱动器上的 CD 盘中读出串行的 MPEG 数据流信息及其他控制信号。经过 MPEG 音频译码器,解出立体声的音频信号再经过模数(A/D)变换器,通过卡拉 OK 处理器可接收话筒输入的卡拉 OK 信号,经过混合叠加处理放大输出到音响设备或电视机。

DVD 播放机的工作原理与 VCD 基本相同,只是视频和音频编码和解码标准采用

MPEG-Ⅱ 或 AC-3。

(3) 多媒体家庭网关(multimedia home gatewake, MHG)

多媒体家庭网关适合家庭应用环境多功能集成,主要功能是:接收并播放数字电视节目、支持多协议的因特网、支持家庭网络控制中心的功能及具有家庭信息服务器的功能。

多媒体家庭网关的硬件结构可分为控制子系统、数字处理子系统、接口子系统、用户/扩展接口子系统。

多媒体家庭网关的软件系统结构可分为设备驱动层、基本操作系统层、逻辑资源层、中间件运行环境层和应用层。

具体结构图见主教材图 1.3。

1.2.4 多媒体技术的应用和发展

1. 多媒体数据库

多媒体数据库和传统的数据库有很大的区别,传统的数据库主要是文字、数据等信息的处理,而多媒体数据库除了文字、数据外,还有图形、图像、声音、动画视频等信息。目前,多媒体数据库还没有较成熟的模型。研究表明,采用面向对象的方法来描述和建立多媒体数据模型是较好的一种方法。面向对象的主要概念包括对象、类、方法、消息、封装和继承集,可以很方便地描述复杂的多媒体信息。

由于多媒体数据,如声音、图像、视频等信息数据量很大,因此,在传输和存储过程中均要进行压缩处理。多媒体数据的存储、管理和存取方法引入了基于内容的检索方法、矢量空间模型信息索引检索技术、超位检索技术及智能索引技术等多种方法。

2. 多媒体通信

多媒体通信可分为两类:

(1) 对称全双工的多媒体通信,如分布式多媒体信息系统、视频会议系统(分为点对点的视频会议系统和多点视频会议系统)以及计算机支撑的协同工作环境。

(2) 非对称全双工的多媒体通信系统,如交互式电视系统(ITV)、点播电视系统(VOD)、远程教育系统、远程医疗诊断系统及远程图书馆等。

多媒体通信要解决两个关键技术问题,即多媒体数据压缩和高速数据通信。

3. 多媒体著作工具和电子出版物

多媒体著作工具可分为基于图符(icon)或流线(line)的创作工具;基于卡片(card)和页面(page)的创作工具;以传统程序语言为基础的创作工具。主要代表产品有 Action, AuthorWare, Icon Auther, ToolBook 以及 Hypercard。多媒体著作工具要具有良好的面向对象的编程环境、具有较强的支撑多媒体数据 I/O 能力。

用多媒体著作工具可以制作各种电子出版物,如清华大学计算机系研制的“金融博士”等多媒体应用系统、演示系统或信息查询系统、培训和教育系统、娱乐、视频动画广告

及专用多媒体应用系统等。

4. 多媒体计算机的发展趋势

(1) 进一步完善计算机支撑的协同工作环境(CSCW)。CSCW 系统是对完成共同任务的群体进行支持,并提出共享环境访问接口的计算机系统。CSCW 系统具有两个本质特征:共同任务和共同环境。一般应具有以下 3 种活动:

- ① 通信。协同工作者之间进行信息交换。
- ② 合作。群体协同共同完成某项任务。
- ③ 协同。对协同工作进行协同,使群体工作和谐,避免冲突和重复。

CSCW 系统可分为 3 种类型:

- ① 交互式,即 CSCW 群体工作者之间的交互可以是同步或异步。
- ② 地理位置,即参与协作的多个用户可以是远程的或本地的。
- ③ 群体规模,即协作既可以是两个人之间的,也可以是多人之间的。

(2) 智能多媒体系统。多媒体计算机从发展看应具有以下智能:

- ① 文字的识别和输入,如印刷体汉字、联机手写体汉字和脱机手写体汉字的识别和输入。
- ② 汉语语音的识别和输入,如特定人、非特定人以及连续汉语语音的识别和输入。
- ③ 自然语言理解和机器翻译,如汉语的自然语言理解和机器翻译、图形的识别和理解、机器人视觉和计算机视觉、知识工程以及人工智能等。

(3) 把多媒体信息实时处理和压缩编码算法集成到 CPU 芯片中。计算机产业的发展趋势应把多媒体和通信的功能集成到 CPU 芯片中,过去的计算机较多考虑计算功能,主要用于数学运算及数值处理,而随着多媒体技术和通信技术的发展,需要计算机能综合处理声、文、图信息,并具备通信的功能。把这些功能和算法集成到 CPU 芯片中,有几项原则,即压缩算法采用国际标准的设计;多媒体功能的单独解决变成集中解决;体系结构设计和算法相结合。

1.3 本章小结

本章对多媒体计算机的定义、分类和多媒体计算机要解决的关键技术,以及多媒体技术的应用和发展,均做了详细的讨论。

多媒体计算机技术是综合处理声、文、图、音频、视频等信息的技术。多媒体计算机具有信息载体的多样性、集成性和交互性。

多媒体计算机的关键技术是解决视频、音频信号的获取和处理,包括多媒体数据的压缩编码和解码技术,以及多媒体数据的输出技术。

多媒体技术促进了通信、娱乐和计算机的融合,为解决电视数字化高清晰度电视提供了切实可行的方案。应用多媒体计算机技术可制作 DVD、影视音响设备,以及制作多媒体家庭网关。多媒体技术的发展促进了多媒体数据库、多媒体通信、多媒体创作工具及应用的发展。多媒体计算机将朝着高分辨率、高速化、简单化、智能化方向发展。

1.4 例题详析

【例题 1】 多媒体计算机中的媒体信息是指_____。

- (1) 数字、文字 (2) 声音、图形 (3) 动画、视频 (4) 以上信息
(A) (1) (B) (2) (C) (3) (D) 全部

答案 (D)

本题考查学生对多媒体信息的理解。多媒体信息应该是以上提到的全部信息,故正确答案是(D)。

【例题 2】 多媒体技术的主要特性有_____。

- (1) 多样性 (2) 集成性 (3) 交互性 (4) 实时性
(A) 仅(1) (B) (1) (2) (C) (1) (2) (3) (D) 全部

答案 (C)

本题考查学生对多媒体技术主要特性的了解。在多媒体技术中,多样性、集成性和交互性是最主要的特性,而实时性不是多媒体技术的主要特性,故正确答案是(C)。

【例题 3】 Commodore 公司于 1985 年在世界上推出的第 1 个多媒体计算机系统是_____。

- (A) Macintosh (B) DVI (C) Amiga (D) CD-I

答案 (C)

本题考查学生对多媒体计算机系统发展历史的了解情况。Commodore 公司于 1985 年推出的第 1 个多媒体系统是 Amiga,所以正确答案是(C)。

1.5 习 题

1. 根据多媒体的特性判断,_____属于多媒体的范畴。

- (1) 交互式视频游戏 (2) 有声图书 (3) 彩色画报 (4) 彩色电视
(A) 仅(1) (B) (1) (2) (C) (1) (2) (3) (D) 全部

2. _____不是多媒体核心软件。

- (1) AVSS (2) AVK (3) DOS (4) AmigaVision
(A) (3) (B) (4) (C) (3) (4) (D) (1) (3)

3. 要把一台普通的计算机变成多媒体计算机,需要解决的关键技术是_____。

- (1) 视频音频信号的获取
(2) 多媒体数据压缩编码和解码技术
(3) 视频音频数据的实时处理和特技
(4) 视频音频数据的输出技术

- (A) (1) (2) (3) (B) (1) (2) (4) (C) (1) (3) (4) (D) 全部

4. Commodore 公司于 1985 年率先在世界上推出了第 1 个多媒体计算机系统 Amiga,其主要功能是_____。

- (1) 用硬件显示移动数据,允许高速的动画制作
 - (2) 显示同步协处理器
 - (3) 控制 25 个通道的 DMA,使 CPU 以最小的开销处理盘、声音和视频信息
 - (4) 从 28Hz 振荡器产生系统时钟
 - (5) 为视频 RAM(VRAM)和扩展 RAM 卡提供所有的控制信号
 - (6) 为 VRAM 和扩展 RAM 提供地址
- (A) (1) (2) (3) (B) (2) (3) (5) (C) (4) (5) (6) (D) 全部

5. 国际标准 MPEG-II 采用了分层的编码体系,提供了 4 种技术,分别是_____。

- (1) 空间可扩展性、信噪比可扩充性、框架技术、等级技术
 - (2) 时间可扩充性、空间可扩展性、硬件扩展技术、软件扩展技术
 - (3) 数据分块技术、空间可扩展性、信噪比可扩充性、框架技术
 - (4) 空间可扩展性、时间可扩充性、信噪比可扩充性、数据分块技术
- (A) (1) (B) (2) (C) (3) (D) (4)

6. 多媒体技术未来发展的方向是_____。

- (1) 高分辨率,提高显示质量
 - (2) 高速度化,缩短处理时间
 - (3) 简单化,便于操作
 - (4) 智能化,提高信息识别能力
- (A) (1)(2)(3) (B) (1)(2)(4) (C) (1)(3)(4) (D) 全部

7. 简述多媒体计算机的关键技术及其应用领域。

第 2 章 音频信息的获取与处理

2.1 本章要点

- (1) 数字化音频的获取与处理的基本概念,模拟音频与数字音频的区别。数字音频采样和量化的基本原理,以及数字音频的文件格式和音频信号的特点。
- (2) 音频卡的工作原理、功能、分类和音频卡的安装使用。
- (3) 音频编码的原理、标准,以及编码解码的基本方法。
- (4) 音乐合成和 MIDI 的接口规范,以及 MIDI 在多媒体技术中的应用,语音识别和合成原理及其分类。

2.2 重点与难点内容分析

2.2.1 数字音频的基本概念

1. 模拟音频与数字音频技术

声音是一种机械波,振动越强,声音就越大。例如,话筒把机械振动转换成电信号,这是一种模拟的音频,它是以模拟电压的幅度表示声音的强弱。

数字音频技术是把表示声音强弱的模拟电压用数字表示,如 0.5V 电压用数字 20 表示,2V 电压用 80 表示。模拟电压的幅度,即使某电平范围内,也可以取无穷多个,如 1.2V,1.21V,1.215V,...。而用数字来表示音频幅度时,只能把无穷多个电压幅度用有限个数字表示。把某一幅度范围的电压用一个数字表示,这叫作量化。

数字音频通过采样量化,把模拟量表示的音频信号转换成许多二进制数 1 和 0 组成的数字音频文件。

2. 数字音频的文件格式与转换

多媒体技术中常用的声音文件格式如下:

(1) WAV 文件

WAV 是 Microsoft 公司的音频文件格式。Microsoft sound system 软件 Sound Finder 可以转换 AIF, SND 和 VOC 文件到 WAV 格式。其中, AIF 是 Apple 计算机的音频文件格式; SND 是 Next 计算机的波形音频文件格式。

(2) VOC 文件

VOC 文件是 Creative 公司波形音频文件格式。

利用声霸卡提供的软件可实现 VOC 和 WAV 文件之间的转换,即程序 VOC2WAV 转换 Creative 的 VOC 文件到 Microsoft 的 WAV 文件;程序 WAV2VOC 转换 Microsoft

的 WAV 文件到 Creative 的 VOC 文件。

(3) MIDI 文件

乐器数字接口(musical instrument digital interface,MIDI)。

RMI 是 Microsoft 公司的 MIDI 文件格式。

2.2.2 音频卡的功能及工作原理

1. 音频卡的功能

音频卡的主要功能有音频录放、编辑、音乐合成、文语转换、CD-ROM 接口、MIDI 接口、游戏接口等。

(1) 音频录放

① 数字化音频采样频率范围：5~44.1kHz；量化位：8 位/16 位；通道数：立体声/单声道。

② 编码与压缩：基本编码方法有 PCM(脉冲编码调制)；压缩编码方法有 ADPCM (8:4, 8:3, 8:2, 16:4)。

CCITT A 一律(13:8)。

CCITT μ 一律(14:8)。

实时硬件压缩和软件压缩。

③ 音频录放的自动动态滤波。

④ 录音声源：麦克风、立体声线路输入、CD。

(2) 编辑与音乐合成

编辑与合成就像一部数字音频编辑器，它可以对声音文件进行各种特殊的处理，如倒播、增加回音效果、静噪音、往返放音、交换声道等。

音乐合成功能和性能主要依赖于合成芯片。目前，Yamaha 的合成器芯片占有率最高，其中主要是 FMOPL 系列。

(3) 其他接口

① MIDI 接口：乐器数字接口的标准，它规定了电子乐器与计算机之间数据通信的协议。

② CD-ROM 接口：目前，音频卡的 CD-ROM 接口有多种，如 Sound Blaster 专用 CD-ROM 接口等。

③ 游戏棒接口：标准的 PC 游戏棒接口，可接一个或两个游戏棒。

(4) 文语转换和语音识别

① 文语转换就是把计算机内的文本文件转换成声音文件。一般音频卡都提供英语文语转换软件，如 Sound Blaster。

② 语音识别软件。有的音频卡提供语音识别软件，如 Sound system 卡上的 Voice pilot 软件，通过这个软件可以利用语音来控制计算机或执行 Windows 下的命令。

2. 音频卡的工作原理

音频卡的工作原理框图见主教材图 2.1。主要由以下几个部分组成。

(1) 声音的合成与处理：这是音频卡的核心部分，它由数字声音处理器、调频(FM)音乐合成器及乐音数字接口(MIDI)控制器组成。它的主要任务是完成声波信号的模数(A/D)和数模(D/A)转换，利用调频技术控制声音的音调、音色和幅度等。

(2) 混合信号处理器：混合信号处理器内置数字模拟混音器，混音器的声源可以是MIDI信号、CD音频、线路输入、麦克风等。可以选择一个声源或几个不同的声源进行混合录音。

(3) 功率放大器：由于混合信号处理器输出的信号功率还不够大，不能推动扬声器或音箱，所以一般都要有一个功率放大器作为功率放大，使得输出的音频信号有足够的功率。

(4) 总线接口和控制器：总线接口有多种，早期的音频卡为ISA总线接口，现在的音频卡一般是PCI总线接口。总线接口和控制器由数据总线双向驱动器、总线接口控制逻辑、总线中断逻辑及直接存储器访问(DMA)控制逻辑组成。

2.2.3 数字音频信号编码算法和标准

音频信号通常采用波形编码。波形编码方法要求重建语音信号：尽可能保持原语音信号的情况。波形编码的对象是语音的波形，算法简单，易于实现，而且对语音的恢复能保持原有语音的特点。但波形编码也有其不足的地方，如易受量化噪声的干扰，进一步降低编码比特数(编码率)也比较困难。常用有如下3种波形编码方法。

(1) 脉冲编码调制(PCM)：直接对语音信号进行模数转换。只要采样频率足够高，量化位数足够多，就能使解码后恢复的语音信号有较高的质量。

(2) 差分脉冲编码调制(DPCM)：即只传输语音预测值和样本值的差值，以此降低音频数据的编码率。

(3) 自适应差分编码调制(ADPCM)：它是DPCM方法的改进，通过调整量化步长，对不同数模设置不同的量化步长，使数据得到进一步的压缩。

还有模型参数编码方法和基于人的听觉特性的编码方法。有关编码方法和国际标准见主教材表2.3。

2.2.4 音乐合成技术——MIDI

1. MIDI 的基本概念

在多媒体技术中，合成音乐或声响效果的方法可以直接采用波形声音产生，但最常用的方法是采用电子乐器数字接口(MIDI)。它不是像波形声音那样的数字化过程，而是将电子乐器键盘的弹奏过程记录下来。例如，将按哪个键、压力多大、时间多长等参数记录下来，作为某一个乐谱的一种数字描述，即MIDI消息。当需要重新产生这个乐谱时，只需要从MIDI文件中读出相应的MIDI消息，再生成对应的乐器声音波形，经放大处理后可就输出。MIDI声音处理过程如图2.1所示。

采用MIDI的最主要优点如下：

(1) 占用数据量小。例如，30分钟的立体声音乐使用CD-DA格式波形存储约需要

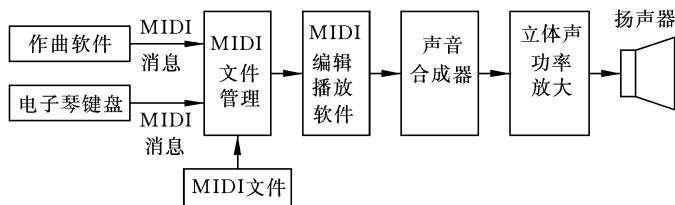


图 2.1 MIDI 声音处理过程

300MB 的存储量,而采用 MIDI 记录时,只需要 200KB,相差 1500 倍,即使波形声音采用 ADPCM 编码也要差两个数量级以上。

(2) 声音配置方便。例如,当多媒体系统中播放波形声音时(如图片的一段解说词),此时若还需要配上某种音乐作为解说的效果,就不可能同时调用两个波形声音文件,而用 Windows 下播放 MIDI 文件记录下来的音乐就很方便。

(3) MIDI 声音编辑修改时方便灵活。可以随意修改曲子的速度、音调,也可以改换乐器的种类,从而产生合适的音乐。

MIDI 的不足之处主要是合成后输出的声音质量取决于 MIDI 硬件,普通多媒体计算机配置的声音卡合成器只适合打击乐器等电子乐器,手风琴、小提琴之类的乐器声音则需要专门配置 MIDI 声音合成器或外接一些高质量的 MIDI 电子乐器。

2. MIDI 接口规范

(1) MIDI 规范规定,每种 MIDI 装置由一个接收器和一个发送器组成。发送器生成符合 MIDI 格式的消息并向外发送,接收器接收 MIDI 格式的消息,并执行 MIDI 命令。MIDI 收发器可用一种通用的异步收发器互相连接。数据传输速率为 31250b/s,每个数据位前后各有一个起始位和停止位。

(2) MIDI 设备有 3 种端口。MIDI 输入口(MIDI-in)用来接收从其他 MIDI 设备发过来的消息;MIDI 输出口(MIDI-out)用来发送本设备产生的原始 MIDI 消息;MIDI 转送口(MIDI-Thru)用来在 MIDI 设备之间进行消息转送。MIDI 设备可同时具有 3 种端口或两种端口,但至少应具备其中一种端口。

(3) MIDI 规范规定,MIDI 键盘为 128 键(比标准 88 键钢琴多 21 个低音符和 19 个高音符),编号为 0~127。MIDI 消息可以描述每个音符的信息,包括对应的键号、按键的持续时间、音量和力度。

(4) MIDI 接收器中有 16 个声道,它们可以同时向声音合成器传送 16 路不同的声音,好像指挥 16 个乐器演奏一样。

(5) MIDI 文件中包含了一连串的 MIDI 消息,每个 MIDI 消息由若干字节组成,通常第 1 个字节为状态字节,其后则为一个或两个数据字节。状态字节的特征是最高位为 1,用来指出紧随其后的数据字节的用途和含义;数据字节的特征是最高位为 0,表示它们是一条 MIDI 消息的信息内容。